Ain Shams Engineering Journal (2016) 7, 653–662

Ain Shams University

Ain Shams Engineering Journal

www.elsevier.com/locate/asej
www.sciencedirect.com

CIVIL ENGINEERING

Using automatic calibration method for optimizing the performance of Pedotransfer functions of saturated hydraulic conductivity

Ahmed M. Abdelbaki *

Fayoum University, Faculty of Engineering, Fayoum, Egypt
Biological and Agricultural Engineering Dept., North Carolina State University, USA

Received 4 November 2014; revised 5 May 2015; accepted 27 May 2015
Available online 3 July 2015

KEYWORDS

Automatic calibration;
Pedotransfer functions;
Saturated hydraulic conductivity;
Shuffled complex evolution

Abstract Pedotransfer functions (PTFs) are an easy way to predict saturated hydraulic conductivity (K_{sat}) without measurements. This study aims to auto calibrate 22 PTFs. The PTFs were divided into three groups according to its input requirements and the shuffled complex evolution algorithm was used in calibration. The results showed great modification in the performance of the functions compared to the original published functions. For group 1 PTFs, the geometric mean error ratio (GMER) and the geometric standard deviation of error ratio (GSDER) values were modified from range (1.27–6.09), (5.2–7.01) to (0.91–1.15), (4.88–5.85) respectively. For group 2 PTFs, the GMER and the GSDER values were modified from (0.3–1.55), (5.9–12.38) to (1.00–1.03), (5.5–5.9) respectively. For group 3 PTFs, the GMER and the GSDER values were modified from (0.11–2.06), (5.55–16.42) to (0.82–1.01), (5.1–6.17) respectively. The result showed that the automatic calibration is an efficient and accurate method to enhance the performance of the PTFs.

1. Introduction

Saturated hydraulic conductivity K_{sat} is needed for many studies and applications related to irrigation, drainage, water movement and solute transport in the soil. Due to spatial variability of soil hydraulic properties, large numbers of measurements are often required to properly characterize such properties even at the field scale. K_{sat} can be obtained from direct laboratory or field measurements, which become too costly and time consuming when many K_{sat} values are required for large scale applications. Alternatively, K_{sat} can be indirectly estimated in terms of the more widely available soil properties such as particle size distribution, bulk density, porosity, and organic matter content. This alternative involves using what is called pedotransfer functions (PTFs) [1]. The term PTF means transferring the data that we have to the data that we need.

Several studies have been conducted to develop PTFs for predicting K_{sat} [2–10]. Other studies have been conducted to evaluate and compare the performance of these functions

* Tel.: +20 1000977229.
E-mail addresses: ama15@fayoum.edu.eg, amabdelb@ncsu.edu.
Peer review under responsibility of Ain Shams University.

ELSEVIER Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.asej.2015.05.012
2090-4479 © 2015 Faculty of Engineering, Ain Shams University. Production and hosting by Elsevier B.V.

[11–13]. The results of these evaluations showed that the PTFs showed different performance (usually poorer) when applied to soil database different from which is used in its development.

The objective of this study is to use the automatic calibration method to adjust the parameters of the published PTFs of saturated hydraulic conductivity to optimize its performance and enable the hydrologist to apply these functions with any soil conditions even if they are different from what were used in their development.

2. Materials and methods

2.1. Pedotransfer functions

Twenty two Pedotransfer functions were calibrated in this study. Due to the lacking of some PTFs inputs in soil database, the functions were classified into three groups according to their input requirements. Group 1 requires inputs of effective porosity only. Group 2 requires inputs of particle size distribution, porosity and bulk density. Group 3 requires inputs of particle size distribution, bulk density and organic matter content.

2.1.1. Group 1 Pedotransfer functions

This type of PTFs, first developed by Ahuja et al. [14], is an empirical relationship between saturated hydraulic conductivity and effective porosity (Eq. (1)).

$$K_{sat} = c * \varphi_e^m \tag{1}$$

where φ_e is the effective porosity, defined as the difference between the total porosity and field capacity (water content at 33 kPa matric potential), c and m are empirical coefficients. The input requirements of these PTFs are available in most soil databases. Six functions of this type were calibrated in this study (Table 1).

Forrest et al. [15] measured K_{sat} in the laboratory for 118 undisturbed soil cores from Australia and used these measurements to obtain the c and m coefficients (F1). Minasny and McBratney [12] used 462 samples from Australian soils to develop PTF for predicting K_{sat} with an effective porosity based on a field capacity determined at 10 kPa (F2). Suleiman and Ritchie [16] used data from 11 homogeneous textural-class soils and several international and American soils to obtain the c and m coefficients of Ahuja et al.'s model (F3). Spychalski et al. [17] developed three PTFs for predicting K_{sat} in terms of the effective porosity. They used a data set of 35 measured soil samples and considered the field capacity as the water content at 10 kPa matric potential. One of these functions (F4) is in the form of Ahuja et al.'s model and the other two PTFs (F5, F6) are multi-linear regressions relating the hydraulic conductivity to the effective porosity.

2.1.2. Group 2 Pedotransfer functions

This group of PTFs requires more inputs than group 1. The PTFs of this group predict K_{sat} in terms of particle size distribution data (%sand, %silt, and %clay), bulk density, and total porosity. Seven PTFs were identified under this category (Table 2).

Puckett et al. [18] used 42 US soil samples to develop an exponential function for predicting K_{sat} in terms of clay content (F7). Using a larger data set of 577 US samples, Dane and Puckett [19] modified the PTF developed by Puckett et al. [18] (F8). Julia et al. [7] used 2178 measured samples from Spain to develop a PTF for predicting K_{sat} in terms of sand content (F9). Cosby et al. [20] used multi-linear regression to develop a PTF using measured soil properties of 1448 US soil samples with inputs of sand and clay contents (F10). Saxton et al. [21] used 230 US data points to derive a PTF for predicting K_{sat} in terms of sand and clay contents (F11). With the same database, Brakensiek et al. [2] developed a PTF for predicting K_{sat} in terms of clay and sand contents and soil porosity (F12). Finally, Jabro [22] used 350 measured samples from international soils to develop PTF for predicting K_{sat} with inputs of silt and clay contents and bulk density (F13).

2.1.3. Group 3 Pedotransfer functions

This group of PTFs requires inputs of particle size distribution, bulk density, and organic matter content. The lack of organic matter measurements in soil databases limits the applicability of the PTFs of this group. Nine PTFs were identified under this category (Table 3).

Julia et al. [7] used 2178 measured samples from Spanish soils to develop a PTF for predicting K_{sat} in terms of sand, clay and organic matter contents (F14). Wösten et al. [4] developed two PTFs using 88 soil profiles from the Netherlands, one PTF for sandy soil (F15) and the other one (F16) for clayey soils. Wösten et al. [5] used 1136 soil samples from the HYdraulic PRoperties of European Soils database (HYPRES) to develop a PTF for predicting K_{sat} in terms of clay, silt and organic matter contents and bulk density (F17). Then, Wösten et al. [6] developed two PTFs for Dutch soils, one function for sandy soils that predicts K_{sat} in terms of silt and organic matter contents and bulk density (F18), and the other function for loamy and clayey soils with inputs of clay and organic matter contents and bulk density (F19). Vereecken et al. [3] used 182 measured soil samples from Belgium to develop a PTF in terms of sand, clay, and organic matter contents and bulk density (F20). Recently, Weynants et al. [23] used the same data set to develop a PTF for predicting K_{sat} in terms of sand and

Table 1 Basic information about original published Pedotransfer functions of group 1.

PTF	Formula (cm/h)	Data set		References
		Size	Source	
F1	$K_{sat} = 0.1 * \text{Exp}(10.8731 + 3.914\ln(\varphi_e))$	118	Australia	[15]
F2	$K_{sat} = 2319.055\varphi_e^{3.66}$	462	Australia	[12]
F3	$K_{sat} = 467.5\varphi_e^{3.15}$	60	International	[16]
F4	$K_{sat} = 4031.57\varphi_e^{3.295}$	35	USA	[17]
F5	$K_{sat} = -2.52 + 581.598\varphi_e^{1.5} - 6966.14\varphi_e^{2.5} + 11693.78\varphi_e^{3}$	35	USA	[17]
F6	$K_{sat} = -3.51 - 18154.6\varphi_e^{1.5} - 12213.8\varphi_e^{2}\ln\varphi_e - 6925.78\varphi_e/\ln\varphi_e$	35	USA	[17]

Table 2 Basic information about original published Pedotransfer functions of group 2.

PTF	Formula (cm/h)[a]	Data set		References
		Size	Source	
F7	$K_{sat} = 15.696\ \mathrm{Exp}(-0.1975\mathrm{CL})$	42	USA	[18]
F8	$K_{sat} = 30.384\ \mathrm{Exp}(-0.144\mathrm{CL})$	577	USA	[19]
F9	$K_{sat} = 0.0920e^{0.0492\mathrm{SA}}$	2178	Spain	[7]
F10	$K_{sat} = 2.54 * 10^{(-0.6\ +\ 0.012\mathrm{SA}\ -\ 0.0064\mathrm{CL})}$	1448	USA	[20]
F11	$K_{sat} = \mathrm{Exp}[12.01 - 0.0755\mathrm{SA} + (-3.895 + 0.03671\mathrm{SA} - 0.1103\mathrm{CL} + 0.00087546\mathrm{CL}^2)/\theta_s]$	230	USA	[21]
F12	$K_{sat} = \mathrm{Exp}[19.524\varphi - 8.9685 - 0.02821\mathrm{CL} + 0.0001811\mathrm{SA}^2 - 0.009413\mathrm{CL}^2 - 8.39522\varphi^2 + 0.07772\mathrm{SA}\varphi - 0.003\mathrm{SA}^2\varphi^2 - 0.01949\mathrm{CL}^2\varphi^2 + 0.000017\mathrm{SA}^2\mathrm{CL} + 0.0273\mathrm{CL}^2\varphi + 0.00143\mathrm{SA}^2 * \varphi - 0.0000035\mathrm{CL}^2\mathrm{SA}]$	230	USA	[2]
F13	$\mathrm{Log}\,K_{sat} = 9.56 - 0.81\mathrm{Log}\,\mathrm{SI} - 1.09\mathrm{Log}\,\mathrm{CL} - 4.64\mathrm{BD}$	350	Intern.	[22]

[a] K_{sat}, saturated hydraulic conductivity (cm/h); SA, sand content (%); SI, silt content (%); CL, clay content (%); BD, bulk density (Mg m^{-3}); φ, total porosity (cm^3/cm^3).

Table 3 Basic information about original published Pedotransfer functions of group 3.

PTF	Formula (cm/h)	Data set		References
		Size	Source	
F14	$K_{sat} = 0.1 * (-4.994 + 0.56728\mathrm{SA} - 0.131\mathrm{CL} - 0.0127\mathrm{OM})$	2178	Spain	[7]
F15	$K_{sat} = 0.04167 * \mathrm{Exp}(9.5 - 1.471 * \mathrm{BD}^2 - 0.688 * \mathrm{OM} + 0.0369 * \mathrm{OM}^2 - 0.332\ln(\mathrm{clay} + \mathrm{silt}))$	88	Netherlands	[4]
F16	$K_{sat} = 0.04167 * \mathrm{Exp}(-43.1 + 64.8\mathrm{BD} - 22.21\mathrm{BD}^2 + 7.02\mathrm{OM} - 0.1562\mathrm{OM}^2 + 0.985\ln(\mathrm{OM}) - 0.01332\mathrm{clay} * \mathrm{OM} - 4.71\mathrm{BD} * \mathrm{OM})$	88	Netherlands	[4]
F17	$K_{sat} = 0.04167 * \mathrm{Exp}[7.755 + 0.035\mathrm{SI} + 0.93 - 0.976\mathrm{BD} - 0.00048\mathrm{CL}^2 - 0.00032\mathrm{SI}^2 + 0.001\mathrm{SI}^{-1} - 0.0748\mathrm{OM}^{-1} - 0.643\ln\mathrm{SI} - 0.0139(\mathrm{BD.CL}) - 0.167(\mathrm{BD.OM}) + 0.298\mathrm{CL} - 0.0331\mathrm{SI}]$	1136	Europe	[5]
F18	$K_{sat} = 0.04167 * \mathrm{Exp}(45.8 - 14.34\mathrm{BD} + 0.001418\mathrm{SI}^2 - 27.5\mathrm{BD}^{-1} - 0.89\ln(\mathrm{SI}) - 0.34\ln(\mathrm{OM}))$	832	Netherlands	[6]
F19	$K_{sat} = 0.04167 * \mathrm{Exp}(-42.6 + 8.71\mathrm{OM} + 61.9\mathrm{BD} - 20.79\mathrm{BD}^2 - 0.2107\mathrm{OM}^2 - 0.0162\mathrm{CL} * \mathrm{OM} - 5.382\mathrm{BD} * \mathrm{OM}$	832	Netherlands	[6]
F20	$K_{sat} = 0.04167 * \mathrm{Exp}[20.62 - 0.96\ln\mathrm{CL} - 0.66\ln\mathrm{SA} - 0.46\ln\mathrm{OM} - 8.43\mathrm{BD}]$	182	Belgium	[3]
F21	$K_{sat} = \mathrm{Exp}(1.9582 + 0.0308\mathrm{SA} - 0.6142\mathrm{BD} - 0.1566\mathrm{OM})/24$	182	Belgium	[23]
F22	$K_{sat} = 0.04167 * \mathrm{Exp}(13.262 - 1.914\ln\mathrm{SA} - 0.974\ln\mathrm{SI} - 0.058\mathrm{CL} - 1.709\ln\mathrm{OM} + 2.885\mathrm{OM} - 8.026\ln\mathrm{BD})$	36	China	[24]

organic matter contents and bulk density (F21). Finally, Li et al. [24] made 36 measurements of K_{sat} for seven soil profiles collected from northern China and developed a PTF for predicting K_{sat} in terms of sand, silt, clay and organic matter contents and bulk density (F22).

2.2. Soil databases

The soil databases used in this study were for U.S. soils. These databases were obtained from different sources. Most of the data were obtained from SOILVISION, an international soil database that contains more than 6000 records of measured soil properties including soil water retention data, K_{sat}, particle size distribution, bulk density, porosity, and organic matter content [25,26]. Another 160 records of K_{sat}, particle size distribution, bulk density, porosity, and organic matter content, were measured for Oklahoma state soils [27]. We also used a database for New Jersey soils, obtained from the Natural Resources Conservation Service's (NRCS) soil database [28]. This database includes measurements of K_{sat}, particle size distribution, porosity, bulk density, organic matter content and soil water retention data. Additionally, we used small data sets found in the literature [19,29]. Many soil records of the databases did not have all the input requirements for all PTFs. Therefore, the soil data obtained from all databases were grouped into three categories according to the input requirements of the three groups of PTFs. First data set was used for evaluating and calibrating the first group of PTFs, includes 1911 records. Second data set was used with group 2 PTFs, includes measurements of K_{sat}, particle size distribution, bulk density (956 samples). The third data set was used with group 3 of PTFs. This data set includes measurements of K_{sat}, particle size distribution, bulk density and organic matter content. The number of records in this data set is 678, which is considerably less than the other two data sets, primarily because of the lack of organic matter content measurements in the soil databases.

2.3. Statistical analysis

The statistical evaluation of the PTFs was conducted using the geometric mean error ratio (GMER, Eq. (2)) and the geometric standard deviation of the error ratio (GSDER, Eq. (3)) [11,12,30].

$$\mathrm{GMER} = \mathrm{Exp}\left(\frac{1}{n}\sum_{i=1}^{n}\ln(\varepsilon_i)\right) \tag{2}$$

$$\mathrm{GSDER} = \mathrm{Exp}\left[\left(\frac{1}{n-1}\sum_{i=1}^{n}[\ln(\varepsilon_i) - \ln(\mathrm{GMER})]^2\right)^{1/2}\right] \tag{3}$$

where n is the number of data points and ε is the error ratio, calculated using Eq. (4).

$$\varepsilon = \frac{K_p}{K_m} \tag{4}$$

where K_p and K_m are predicted and measured hydraulic conductivities, respectively. A GMER of 1 corresponds to an exact match between the measured and predicted values. A less than 1 GMER indicates underestimation and greater than 1 GMER indicates overestimation by the predictive model. A GSDER of 1 corresponds to a perfect match and it grows with the deviation of predictions from the measured values. The value of normalized mean square error, (NRMSE, Eq. (5)) was used to evaluate the error between predictions and measurements.

$$\mathrm{NRMSE} = \frac{\sqrt{\frac{\sum_{i=1}^{n}(K_{pi}-K_{mi})^2}{N}}}{\overline{K_m}} \tag{5}$$

where $\overline{K_m}$ is the average of the measured saturated hydraulic conductivities.

3. Automatic calibration process

3.1. Optimization algorithm (shuffled complex evolution)

The optimization algorithm used in this study is the Shuffled Complex Evolution-University of Arizona (SCE-UA) it is a global search algorithm for the optimization of a single objective function for up to 16 parameters. Detailed description of this algorithm can be found in [31]. In brief, the procedure of this method can be summarized in the following steps. First, SCE-UA selects an initial population, which is the number of parameters sets, by random sampling throughout the feasible parameters space for p parameters to be optimized. The population size is defined by estimating the number of complexes (group of parameter sets) and the number of parameter sets in each complex to be $(2P + 1)$. The second step is to run the calibrated model using these parameter sets and calculate the objective function between the model predictions and the observations for each simulation. Third, the population size is sorted according to the values of the objective function (i.e. the population is ranked from the minimum to the maximum value of the objective function). Fourth, the population is divided into several groups (complexes) in such a way that the ith complex contains every NGS * $(K - 1) + i$ ranked points, where NGS is the number of complexes, $K = 1, 2, \ldots$ NPG, NPG is the number of points in each complex. Fifth, each complex evolves independently using the simplex algorithm [32] by developing new points that has less objective functions. Sixth, the complexes are periodically shuffled to form new complexes in order to share the gained information. Lastly, the algorithm test for the termination criteria, which is the condition to reach the global optimum either it is the value of the objective function or exceed the maximum number of simulations. If the termination criteria were not met, the previous steps will be repeated. This algorithm searches over the whole parameters spaces and finds the global optimum with a success rate of 100% [33].

3.2. Objective function and evaluation criteria

The automatic calibration process has been conducted to the Pedotransfer function of saturated hydraulic conductivity by adjusting the regression coefficients of the equation to get optimum matching between predicted and measured saturated hydraulic conductivities. The regression coefficients were adjusted between lower limit equal to 10% of the current equation coefficients and upper limit equal to ten times the current published values. For each parameter set, the objective function (Eq. (6)) is calculated between the predicted and the measured K_{sat}.

$$\text{Obj Fun.} = \sqrt{(1-\mathrm{GMER})^2 + (1-\mathrm{GSDER})^2} \tag{6}$$

The calibration process is to minimize the value of the objective function. The new form of the calibrated Pedotransfer function is in the same form of the original published function with the regression coefficient corresponding to the minimum objective function.

4. Results and discussions

4.1. Group 1 Pedotransfer functions

The calibrated forms of PTFs of group 1 that require inputs of effective porosity only are presented in Table 4. The forms of the calibrated functions are the same as the original functions with the change of the regression coefficient only. The result of statistical comparison between calibrated and original functions is presented in Table 5. Also, visual comparison between the measured and predicted values of saturated hydraulic conductivities is presented in Fig. 1 for the original functions and Fig. 2 for the calibrated functions. The calibration process has been conducted to adjust the regression coefficients of the PTFs between lower and upper limits ranged from 10% to 1000% of the current coefficients of the original functions. The calibration process chooses different parameter sets between the lower and upper limits randomly and calculates the objective function between the predictions and measurements for thousands of times. The calibrated function is the

Table 4 New formulas of calibrated Pedotransfer functions of group 1.

PTFs	Formula (cm/h)
F1	$K_{sat} = 0.89879 * \mathrm{Exp}(5.2598 + 2.58794\ln(\varphi_e))$
F2	$K_{sat} = 106.8628\varphi_e^{2.014632}$
F3	$K_{sat} = 172.971\varphi_e^{2.587885}$
F4	$K_{sat} = 106.8628\varphi_e^{2.014632}$
F5	$K_{sat} = -5.3734132 + 310.44732\varphi_e^{1.5} - 2373.2878\varphi_e^{2.5} + 2190.6295\varphi_e^{3}$
F6	$K_{sat} = -4.8977523 - 5996.464\varphi_e^{1.5} - 4317.3993\varphi_e^{2}\ln\varphi_e - 2145.7628\varphi_e/\ln\varphi_e$

Table 5 Comparison between original and calibrated PTFs of group 1.

PTFs	Original functions			Calibrated functions		
	GMER	GSDER	NRMSE	GMER	GSDER	NRMSE
F1	5.157	6.291	358.4	1.00	5.85	151.8
F2	1.893	7.011	149.7	1.01	5.77	153.2
F3	1.273	5.929	140.0	1.00	5.85	151.8
F4	5.654	6.556	384.0	1.01	5.77	153.2
F5	6.059	5.266	359.4	0.91	4.88	143.4
F6	6.094	5.552	435.4	1.15	4.97	134.6

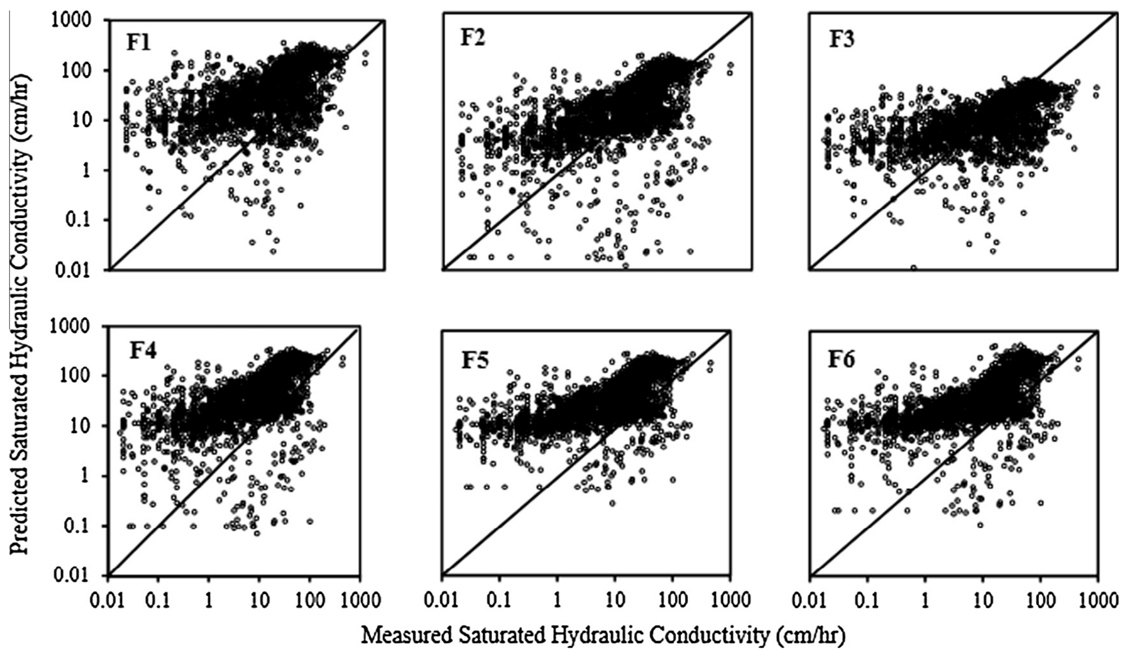


Figure 1 Predicted vs. measured values of K_{sat} for original PTFs of group 1.

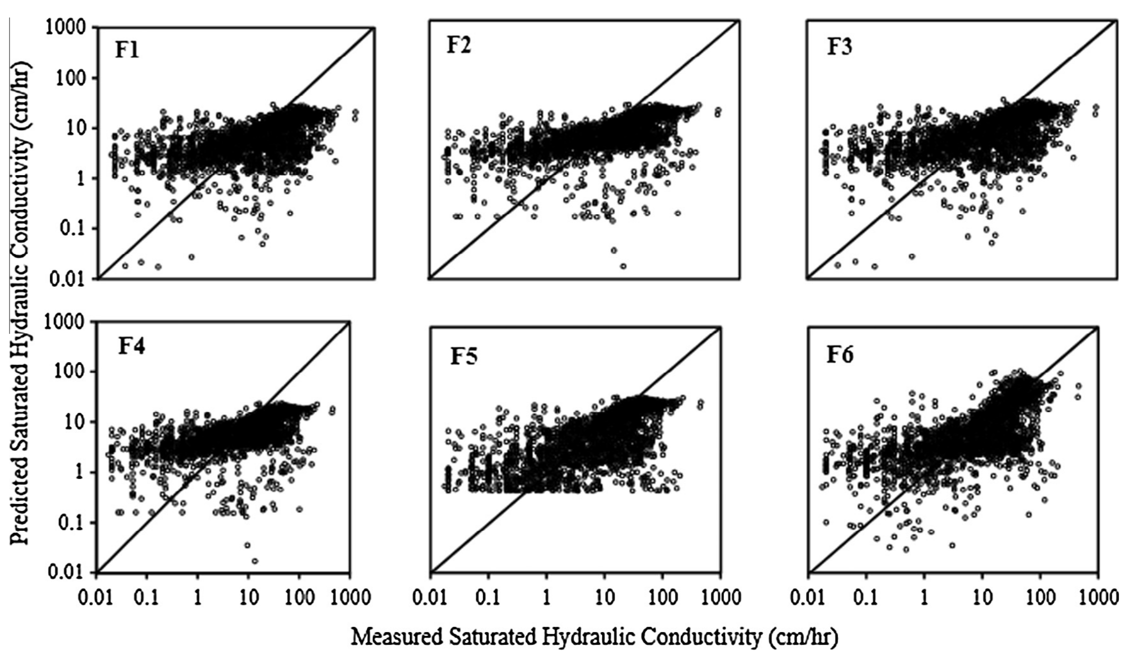


Figure 2 Predicted vs. measured values of K_{sat} for calibrated PTFs of group 1.

Table 6 New formulas of Calibrated Pedotransfer functions of group 2.

PTF	Formula (cm/h)
F7, F8	$K_{sat} = 5.3114\mathrm{Exp}(-0.0479\mathrm{CL})$
F9	$K_{sat} = 0.68596e^{0.022851\mathrm{SA}}$
F10	$K_{sat} = 14.9181 * 10^{(-1.03539+0.00706756\mathrm{SA}-0.00902024\mathrm{CL})}$
F11	$K_{sat} = \mathrm{Exp}[58.19 - 0.357918\mathrm{SA} + (-28.7342 + 0.185766\mathrm{SA} + 0.02694\mathrm{CL} + 0.0003262\mathrm{CL}^2)/\theta_s]$
F12	$K_{sat} = \mathrm{Exp}[9.05198\varphi - 3.46 - 0.04269\mathrm{CL} + 0.00001829\mathrm{SA}^2 - 0.000948507\mathrm{CL}^2 - 0.883269\varphi^2 + 0.040881\mathrm{SA}\varphi - 0.002884\mathrm{SA}^2\varphi^2 - 0.0084376\mathrm{CL}^2\varphi^2 + 0.000001796\mathrm{SA}^2\mathrm{CL} + 0.006906\mathrm{CL}^2\varphi + 0.001229\mathrm{SA}^2 * \varphi - 0.0000014446\mathrm{CL}^2\mathrm{SA}]$
F13	$\mathrm{Log}\,K_{sat} = 2.7876 - 0.2261\mathrm{Log\,SI} - 0.4976\mathrm{Log\,CL} - 1.0764\mathrm{BD}$

Table 7 Comparison between original and calibrated PTFs of group 2.

	Original functions			Calibrated functions		
	GMER	GSDER	NRMSE	GMER	GSDER	NRMSE
F7	0.305	12.383	323.9	1.01	5.93	316.1
F8	1.335	8.309	319.7	1.01	5.93	316.1
F9	0.614	6.768	324.7	1.01	5.85	315.9
F10	0.988	5.953	323.8	1.01	5.78	315.7
F11	0.783	7.264	323.1	1.00	5.76	315.5
F12	0.731	11.622	321.2	1.00	5.54	315.0
F13	1.550	11.710	21034.8	1.03	5.57	313.9

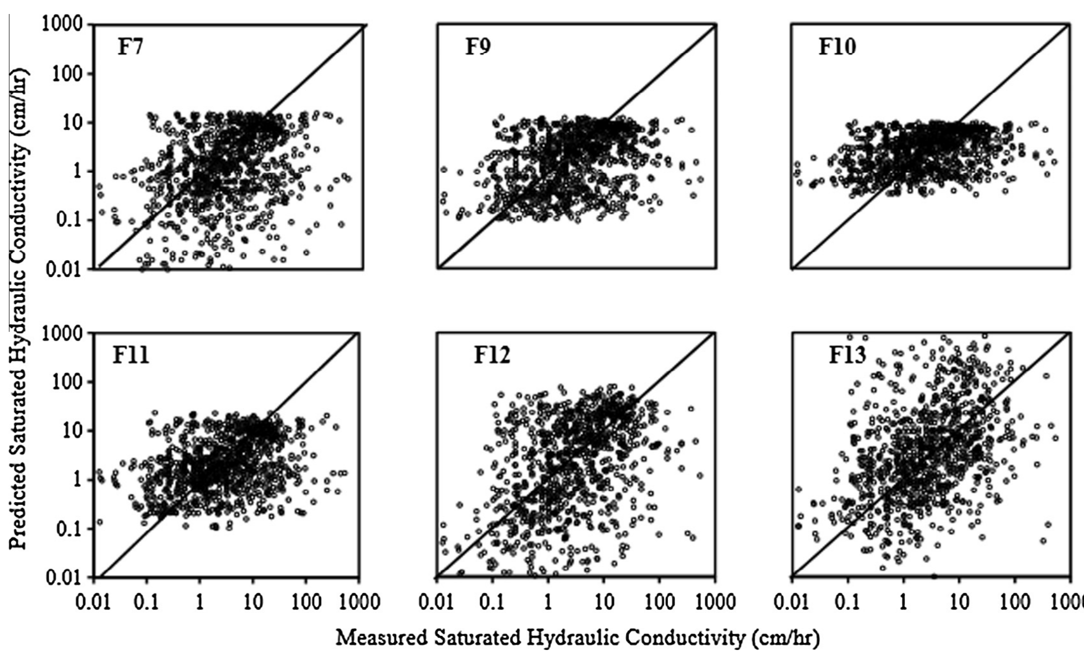


Figure 3 Predicted vs. measured values of K_{sat} for original PTFs of group 2.

function that has the regression coefficients that gave the minimum objective function.

The statistical measures of original and calibrated functions in Table 5 showed great improvements in the performance of all PTFs in terms of the geometric mean of error ratio (GMER), values of geometric standard deviation of error ratio (GSDER) for all PTFs. For the first four functions (F1–F4), the values of GMER are almost 1, which is the perfect value. For F5, GMER is 0.91, which indicates a little bit underestimation while it shows small overestimation 1.15 for F6. The values of GSDER were decreased for all PTFs, which indicate enhancing the predictions for all calibrated functions. The values of NRMSE showed some differences in predictions. It was decreased with the calibration process for all PTFs except for F2 and F3 and it showed small increase in the error in the calibrated functions. The functions that showed great improvements in the values of GMER also showed great improvements in the values of the NRMSE (e.g. F1, F4, F5

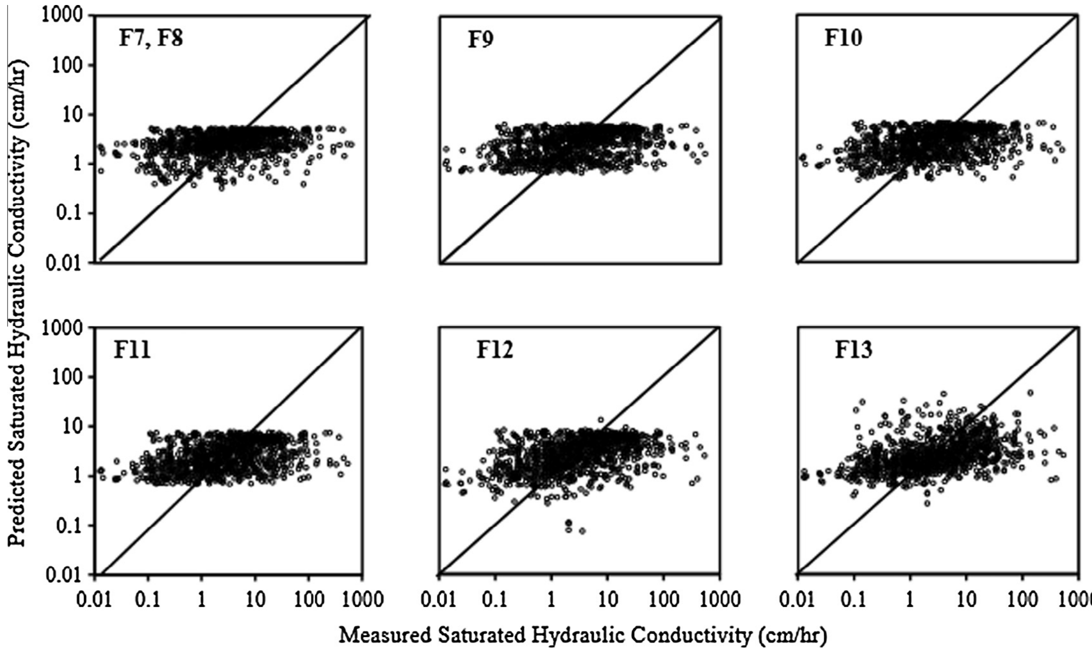


Figure 4 Predicted vs. measured values of K_{sat} for calibrated PTFs of group 2.

and F6) while the PTFs that showed small improvement in the GMER values, did not improve the NRMSE values (e.g. F2 and F3).

In addition to the statistical comparison, the visual comparisons between the measured and predicted values of K_{sat} are presented in Fig. 1 for the original functions and Fig. 2 for the calibrated functions. The results showed visual improvement in the performance of all PTFs. The shape of the scatter diagram in Fig. 1 agrees with the values of statistical measures in Table 4 which showed overestimation for all PTFs. Fig. 2 shows improvement in the performance of the Functions which agree with the statistical measures of GMER values of 1 which means equal distribution of the point around the 1:1 line.

4.2. Group 2 Pedotransfer functions

The calibrated forms of the Pedotransfer functions of group 2 that require inputs of particle size distribution, bulk density and porosity are presented in Table 6. These functions were calibrated using the second data set which contains 956 soil samples. The calibrated functions have the same form as the original functions with only the change in the regression coefficients. The calibrated forms of F7 and F8 have the same functions since it have the same shape of original functions and their calibrated regression coefficients are the same.

Table 7 shows the statistical comparison between the original and calibrated Pedotransfer functions. Five out of the seven PTFs showed underestimation in the original functions (GMER < 1) while the other two functions showed overestimation. After the calibration process, the GMER values of all functions were modified to almost 1 which is the ideal value. Also, the values of the GSDER were modified and decreased for all functions which mean that the performance has been enhanced. The value of the normalized root mean square error (NRMSE) showed a little decrease for all PTFs except for F13 which showed great reduction in the error.

In addition to the statistical comparison between the original and calibrated functions, the visual comparisons are

Table 8 New formulas of calibrated Pedotransfer functions of group 3.

PTF	Formula (cm/h)
F14	$K_{sat} = (-4.2763 + 0.10728\text{SA} - 0.011547\text{CL} - 0.005746\text{OM})$
F15	$K_{sat} = 0.04166667 * \text{EXP}(6.45 - 0.3 * \text{BD}^2 - 0.1 * \text{OM} + 0.0064736 * \text{OM}^2 - 0.40794 * \ln(\text{clay} + \text{silt}))$
F16	$K_{sat} = 0.041667 * \text{EXP}(-11.19 + 21.28\text{BD} - 6.995\text{BD}^2 + 4.46192\text{OM} - 0.145012\text{OM}^2 + 0.2 * \text{LN(OM)} - 0.0634823\text{clay} * \text{OM} - 2.87372 * \text{BD} * \text{OM})$
F17	$K_{sat} = 0.04167 * \text{Exp}[6.06889 + 0.004086\text{SI} + 2.1612 - 1.5415\text{BD} - 0.000099211\text{CL}^2 - 0.0000334\text{SI}^2 + 0.0066789\text{SI}^{-1} - 0.0076953\text{OM}^{-1} - 0.73334\ln\text{SI} - 0.0016279\ (\text{BD.CL}) - 0.018934\ (\text{BD.OM}) + 0.029885\text{CL} - 0.0033359\text{SI}]$
F18	$K_{sat} = 0.041667 * \text{Exp}(11.5106 - 2.826\text{BD} + 0.000184525\text{SI}^2 - 2.92756\text{BD}^{-1} - 0.479674\ln(\text{SI}) - 0.034\ln(\text{OM}))$
F19	$K_{sat} = 0.041667 * \text{Exp}(-6.4783 + 3.46114\text{OM} + 14.5945\text{BD} - 4.88278\text{BD}^2 - 0.132778\text{OM}^2 - 0.0506746\text{CL} * \text{OM} - 1.9866\text{BD} * \text{OM})$
F20	$K_{sat} = 0.041667 * \text{Exp}[7.8461 - 0.92974\ln\text{CL} - 0.0664\ln\text{SA} - 0.12249\ln\text{OM} - 0.92428\text{BD}]$
F21	$K_{sat} = \text{Exp}\ (4.7354 + 0.030636\text{SA} - 1.6039\text{BD} - 0.01566\text{OM})/24$
F22	$K_{sat} = 0.04167 * \text{Exp}(8.5553 - 0.33825\ \ln\text{SA} - 0.78479\ln\text{SI} - 0.035748\text{CL} - 0.171\ln\text{OM} + 0.49858\text{OM} - 1.7049\ln\text{BD})$

Table 9 Comparison between the performances of original and calibrated PTFs of group 3.

PTFs	Original functions			Calibrated functions		
	GMER	GSDER	NRMSE	GMER	GSDER	NRMSE
F14	0.803	6.034	197.3	0.82	5.66	194.9
F15	1.542	8.127	209.8	1.00	5.31	185.5
F16	0.11	16.42	487.1	1.00	6.47	196.6
F17	1.430	6.090	190.2	1.00	5.79	183.5
F18	2.058	9.653	370.9	1.00	5.69	194.9
F19	0.328	10.960	250.5	1.00	6.17	197.3
F20	0.659	9.262	211.7	1.00	5.42	191.5
F21	0.260	5.555	205.5	1.00	5.42	192.3
F22	0.201	8.280	224.6	1.01	5.10	189.7

presented in Fig. 3 for the original functions and Fig. 4 for the calibrated functions. The scatter diagrams showed better performance for the calibrated functions especially for some functions like F7, F12 and F13. For these functions the great reductions in the GSDER values form 12 to 5.5 returns in great modification in the deviation of the predictions from the measured values and enhance the shape of the scatter diagrams.

4.3. Group 3 Pedotransfer functions

Results of PTFs calibration are presented in Table 8. The PTFs of this group require inputs of particle size distribution, bulk density and organic matter content. The calibration process has been conducted using the soil data set of group 3 which contains smaller number of records due to the lacking of the organic matter content in the soil survey. The forms of the calibrated function have the same structure as the original published functions with the only changes are the regression coefficients. These coefficients were adjusted between lower limit of 10% and upper limit of 1000% of the current values in the original functions.

The results of statistical comparison between the original and calibrated functions are presented in Table 9. Six out of the nine functions in its original format underestimate the saturated hydraulic conductivities with GMER values ranged from 0.11 to 0.803 while the other three functions showed

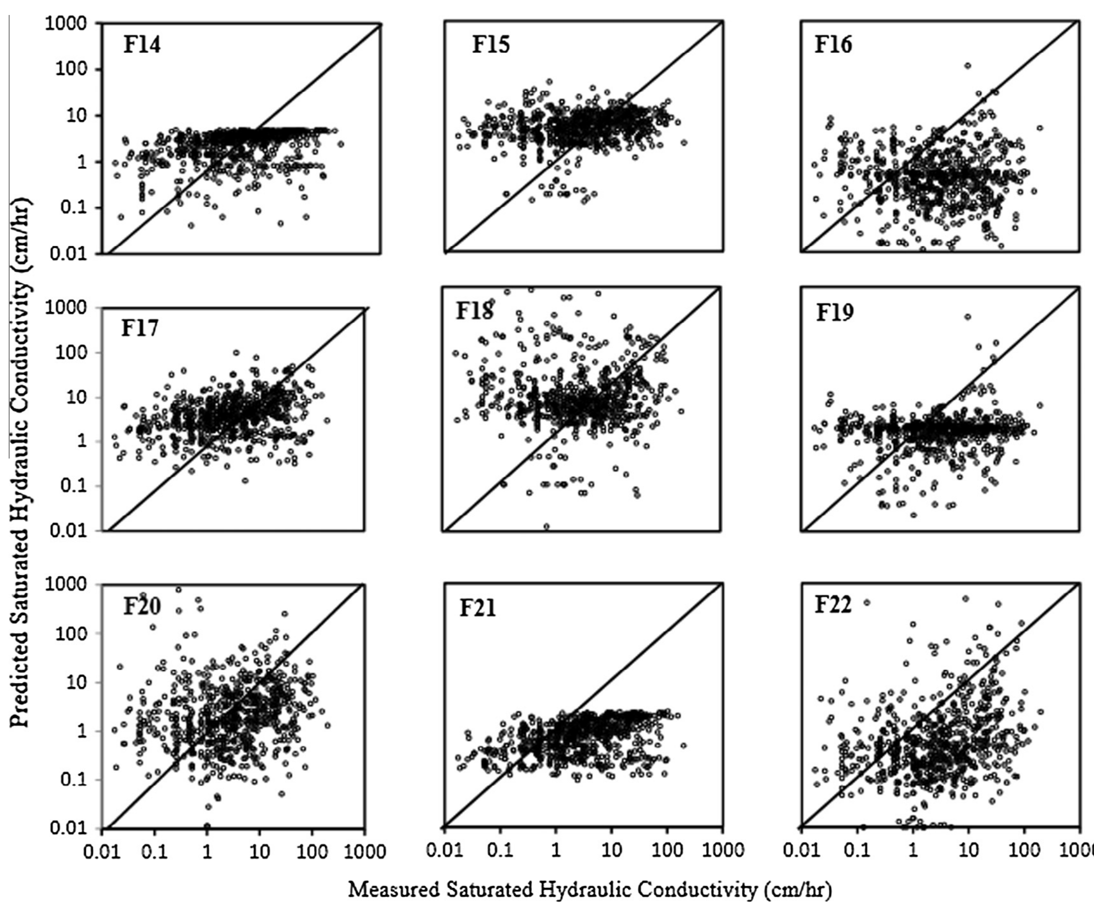


Figure 5 Predicted vs. measured values of K_{sat} for original PTFs of group 3.

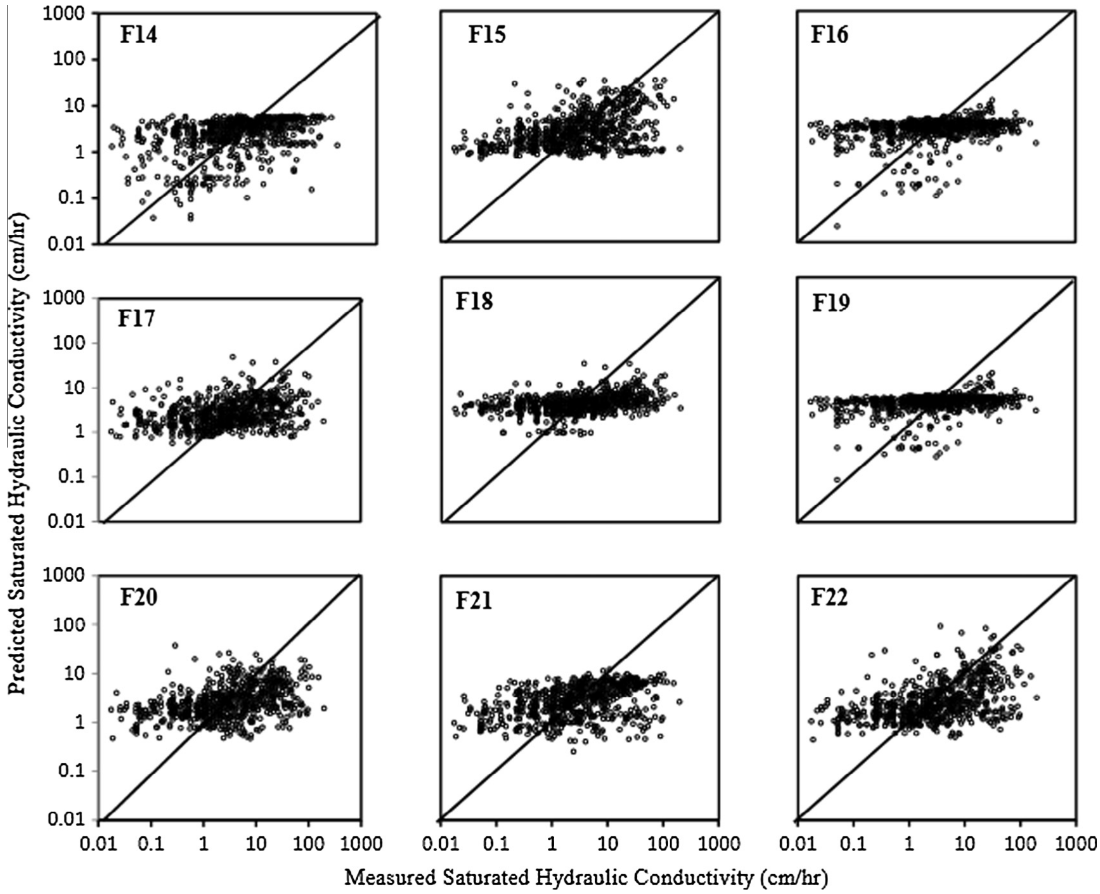


Figure 6 Predicted vs. measured values of K_{sat} for calibrated PTFs of group 3.

overestimation with GMER values of 1.43–2.06. The GSDER values for original functions were from 5.55 to 16.42. The calibration process enhanced the performance of all PTFs to have GMER value of 1 for all PTFs except for F14 (GMER = 0.82). The GMER values of the calibrated functions represent the optimum performance with the average ratio between the predicted and measured values are 1. Also, the values of the standard deviation of error ratio (GSDER) were decreased for all PTFs with the range of 5.1–6.47. The values of the normalized root mean square error were decreased for all PTFs with values ranged from low (i.e. F14, F17 and F21) to high (i.e. F19, F18 and F16). These results showed the great impact of the automatic calibration process in enhancing the performance of the PTFs.

In addition to the statistical comparison between the original published Pedotransfer functions and the calibrated functions, the visual comparison has been conducted. Figs. 5 and 6 show the predicted vs measured saturated hydraulic conductivities values for both original and calibrated functions. Fig. 5 shows the bad performance of the original functions which is clear from the wide spread of the data point around the domain due to the large value of the standard deviation and root mean square error, especially for the functions (F16, F18, F19, F20 and F22). The spread of the data point has been decreased in Fig. 6 and the point was close to the 1:1 due to the reduction in the error and low values of the standard deviation.

5. Conclusion

The Pedotransfer functions of saturated hydraulic conductivity in its current form show high error ratio in predictions. This error returns to the change in the soil conditions of the application database from the development database. This study was conducted to overcome this shortage by modifying the current forms of the published functions using the automatic calibration process by adjusting the regression coefficients of the equation in order to get the optimum match between function predictions and measurements. Twenty-two PTFs were calibrated in this study. These functions were grouped into three groups according to the availability of its inputs. Group 1 requires inputs of effective porosity only, group 2 requires inputs of particle size distribution, bulk density and porosity, and group 3 requires organic matter content in addition to the input requirements of group 2. The calibrated functions showed much better performance than the original functions. For most of the functions, the values of the GMER values were close to 1 which is the optimum value for the ratio between the predictions and measurements. Also, the values of the GSDER and NRMSE showed great reductions in all calibrated functions. These results indicate that the automatic calibration is an efficient method to optimize the performance of the current PTFs when applied in different soil conditions.

References

[1] Bouma J. Using soil survey data for quantitative land evaluation. Adv Soil Sci 1989;9:77–213.

[2] Brakensiek DL, Rawls WJ, Stephenson GR. Modifying SCS hydrologic soil groups and curve numbers for rangeland soils. ASAE paper no. PNR-84203. St. Joseph, MI; 1984.

[3] Vereecken H, Maes J, Feyen J. Estimating unsaturated hydraulic conductivity from easily measured soil properties. Soil Sci 1990;149:1–12.

[4] Wösten JHM. Pedotransfer functions to evaluate soil quality. In: Gregorich EG, Carter MR, editors. Soil quality for crop production and ecosystem health, vol. 5. Amsterdam: Developments in Soils Science Elsevier; 1997. p. 221–45.

[5] Wösten JHM, Lilly A, Nemes A, Bas CL. Development and use of a database of hydraulic properties of European soils. Geoderma 1999;90:169–85.

[6] Wösten JHM, Pachepsky YA, Rawls WJ. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. J Hydrol 2001;251:123–50.

[7] Julia MF, Monreal TE, Jimeneza ASC, Garc Melendez E. Constructing a saturated hydraulic conductivity map of Spain using Pedotransfer functions and spatial prediction. Geoderma 2004;123:257–77.

[8] Nemes A, Rawls WJ, Pachepsky YA. Influence of organic matter on the estimation of saturated hydraulic conductivity. Soil Sci Soc Am J 2005;69:1330–7.

[9] Saxton KE, Rawls WJ. Soil water characteristic estimates by texture and organic matter for hydrologic solutions. Soil Sci Soc Am J 2006;70:1569–78.

[10] Guarracino L. Estimation of saturated hydraulic conductivity Ks from the van Genuchten shape parameter α. Water Resour Res 2007;43:W11502. http://dx.doi.org/10.1029/2006WR005766.

[11] Tietje O, Hennings V. Accuracy of the saturated hydraulic conductivity prediction by Pedotransfer functions compared to the variability within FAO textural classes. Geoderma 1996;69:71–84.

[12] Minasny B, McBratney AB. Evaluation and development of hydraulic conductivity Pedotransfer functions for Australian soil. Aust. J. Soil Res. 2000;38:905–26.

[13] Sobieraj JA, Elsenbeer H, Vertessy RA. Pedotransfer functions for estimating saturated hydraulic conductivity: implications for modeling storm flow generation. J Hydrol 2001;251(3/4):202–20.

[14] Ahuja LR, Naney JW, Green RE, Nielsen DR. Macroporosity to characterize spatial variability of hydraulic conductivity and effects of land management. Soil Sci Soc Am J 1984;48(4):699–702.

[15] Forrest JA, Beatty J, Hignett CT, Pickering JH, Williams RGP. A survey of the physical properties of Wheatland soils in eastern Australia. CSIRO Australia Division of Soils, Divisional Report No. 78; 1985.

[16] Suleiman AA, Ritchie JT. Estimating saturated hydraulic conductivity from soil porosity. Soil Sci Soc Am J 2000;44(2):235–339.

[17] Spychalski M, Kaźmierowski C, Kaczmarek Z. Estimation of saturated hydraulic conductivity on the basis of drainage porosity. Electr J Pol Agr Univ 2007;10(1):#04.

[18] Puckett WE, Dane JH, Hajek BF. Physical and mineralogical data to determine soil hydraulic properties. Soil Sci. Soc. Am. J. 1984;49:831–6.

[19] Dane JH, Puckett W. Field soil hydraulic properties based on physical and mineralogical information. In: van Genuchten MT et al., editors. Proceedings of the international workshop on indirect methods for estimating the hydraulic properties of unsaturated soils. Riverside (CA): University of California; 1994. p. 389–403.

[20] Cosby BJ, Hornberger GM, Clapp RB, Ginn TR. A statistical exploration of soil moisture characteristics to the physical properties of soils. Water Resour Res 1984;20:682–90.

[21] Saxton KE, Rawls WJ, Romberger JS, Papendick RI. Estimating generalized soil water characteristics from texture. Soil Sci Soc Am J 1986;50:1031–6.

[22] Jabro JD. Estimation of saturated hydraulic conductivity of soils from particle size distribution and bulk density data. Trans ASAE 1992;35(2):557–60.

[23] Weynants M, Vereecken H, Javaux M. Revisiting Vereecken Pedotransfer functions: Introducing a closed-form hydraulic model. Vadose Zone J. 2009;8:86–95.

[24] Li Y, Chen D, White RE, Zhu A, Zhang J. Estimating soil hydraulic properties of Fengqiu County soils in the North China Plain using pedo-transfer functions. Geoderma 2007;138:261–71.

[25] Fredlund MD, Sillers WS, Fredlund DG, Wilson GW. Design of a knowledge-based system for unsaturated soil properties. In: 3rd Canadian conference on computing in civil and building engineering, 26–28 August, 1996, Montreal, Quebec.

[26] Fredlund MD, Wilson GW, Fredlund DG. Estimation of hydraulic properties of an unsaturated soil using a knowledge-based system. In: 2nd international conference on unsaturated soils, 27–30 August, 1998, UNSAT'98 Beijing, China.

[27] Mohanty BP, Shouse PJ, Miller DA, van Genuchten MTh. Soil property database: Southern great plains 1997 hydrology experiment. Water Resour Res 2002;38(5):1047.

[28] National Cooperative Soil Survey. Soil Survey Laboratory data and descriptions for some soils of New Jersey. New Brunswick (NJ): Soil Conservation Services, USDA and New Jersey Agric. Exp. Sta., Rutgers Univ; 1974.

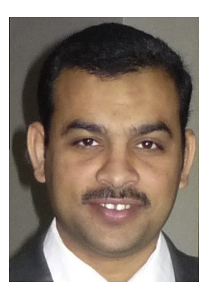
[29] Rawls W, Gimenez JD, Grossman R. Use of soil texture, bulk density, and slope of the water retention curve to predict saturated hydraulic conductivity. Trans ASAE 1998;41(4):983–8.

[30] Wagner B, Tarnawski VR, Hennings V, Müller U, Wessoleke G, Plagge R. Evaluation of Pedotransfer functions for unsaturated soil hydraulic conductivity using an independent data set. Geoderma 2001;102:275–97.

[31] Duan QY, Sorooshian S, Gupta V. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour Res 1992;28(4):1015–31.

[32] Nelder JA, Mead R. A simplex method for function minimization. Comput J 1965;7:308–13.

[33] Sorooshian S, Duan Q, Gupta VK. Calibration of rainfall-runoff models: application of global optimization to the sacramento soil moisture accounting model. Water Resour Res 1993;29:1185–94.



Ahmed M. Abdelbaki is currently a lecturer in the Department of Civil Engineering, Fayoum University, Egypt. He received his Ph.D. degree in Irrigation and Hydraulics from Cairo University, Egypt, and North Carolina State University (NCSU), USA. He studied his Ph.D. in NCSU, USA, from 2008 to 2011 and Postdoctoral in 2014 as visiting scholar. His research interest is agriculture drainage using DRAINMOD model and predicting soil properties using Pedotransfer Functions. Also, his research interests include development of Automatic Calibration algorithms to calibrate hydrologic models.